FINGERPRINT IMAGE SEGMENTATION USING DATA MANIFOLD CHARACTERISTIC FEATURES

ANTÓNIO R. C. PAIVA* and TOLGA TASDIZEN*,†

* Scientific Computing and Imaging Institute,
† Electrical and Computer Engineering Department,
University of Utah, Salt Lake City, UT 84112, USA
arpaiva@sci.utah.edu

Automatic fingerprint identification systems (AFIS) have been studied extensively and are widely used for biometric identification. Given its importance, many well engineered methods have been developed for the different stages that encompass those systems. The first stage of any such system is the segmentation of the actual fingerprint region from the background. This is typically achieved by classifying pixels, or blocks of pixels, based on a set of features. In this paper, we describe novel features for fingerprint segmentation that express the underlying manifold topology associated with image patches in a local neighborhood. It is shown that fingerprint patches seen in a high-dimensional space form a simple and highly regular circular manifold. The characterization of the manifold topology suggests a set of optimal features that characterize the local properties of the fingerprint. Thus, fingerprint segmentation can be formulated as a classification problem based on the deviation from the expected topology. This leads to features that are more robust to changes in contrast than mean, variance and coherence. The superior performance of the proposed features for fingerprint segmentation is shown in the eight datasets from the 2002 and 2004 Fingerprint Verification Competitions.

Keywords: fingerprint segmentation; manifold characterization; feature extraction; dimensionality reduction.

1. Introduction

Fingerprints are widely used biomarkers. Their uniqueness and near immutability over life have been known for more than a century [31]. Because of their initial importance in forensics and the recent need for automated biometric identification systems, automatic fingerprint identification systems (AFIS) have been studied extensively and are still an highly active research topic [25, 7, 47, 37]. As a result, many well engineered methods have been proposed for the different stages that encompass such a system (Fig. 1).

The first stage of a fingerprint identification system is the segmentation of the actual fingerprint region from the background. This pre-processing step allows the subsequent steps to focus only on the actual fingerprint region, which typically
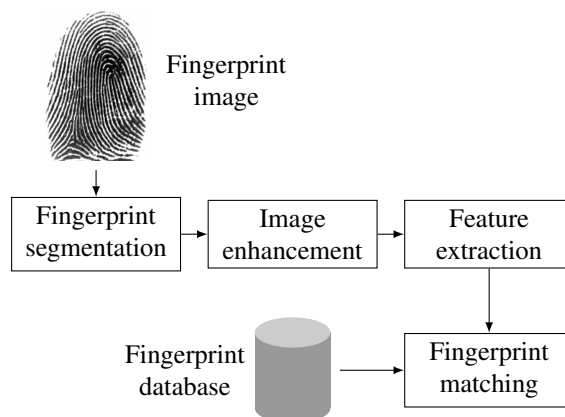
2   *Antonio R. C. Paiva and Tolga Tasdizen*

Fig. 1. Diagram of the stages of a fingerprint matching system.

leads to savings in computation time. More importantly, it prevents extraction of spurious features in the background because of noise. The next stage is image enhancement, aimed at increasing the contrast and reducing noise in preparation for feature extraction. The most commonly used approach for this stage is to filter the fingerprint image with a directional Gabor filter using parameters chosen to match the ridge/valley pattern of fingerprints [18]. Then, a set of features is extracted for matching. The most relevant features are the ridges and minutiae. The ridge features are generally used for fingerprint alignment [21] and for minutia detection [31]. Ultimately, the minutia points are used for fingerprint matching or identification because the relative location of these points is unique to each fingerprint.

Fingerprint segmentation has received considerable attention in the literature. This is understandable since this is the first stage of an AFIS, and any errors at this stage will propagate to the following stages and inevitably decrease the matching performance. Typically, fingerprint segmentation is achieved by classifying pixels, or blocks of pixels, based on a set of features which quantifies an intuitive understanding of how fingerprint regions appear [31, 5, 50]. Although these methods achieve good results for high-quality / high-contrast fingerprints, their segmentation performance is not consistent for low quality fingerprints.

In this paper we describe how a fingerprint can be described in terms of the manifold topology of local patches and the implications of this for fingerprint segmentation. We derive features based on the analysis of the intrinsic structure associated with the fingerprint pattern. This "intrinsic structure" is contained in the manifold topology inherent from local image neighborhoods (i.e., "patches"), and can be expressed in a low-dimensional space which reveals this structure. More specifically, it is shown that fingerprint patches form locally a simple and highly regular circular manifold. This observation suggests a natural set of features that describe a fingerprint region. A key advantage is that the manifold topology derived

from the fingerprint image is a more robust descriptor since it characterizes intrinsic properties that invariant to translation, rotation and contrast. These features are shown to yield better and more robust fingerprint segmentation results in images from several datasets.

2. Background and previous work

Roughly speaking, there are two general approaches to fingerprint segmentation, pixel-wise or block-wise, with the latter being the prominent. The methodology is fundamentally the same, but the computation of the features, at every pixel versus at every block, gives rise to significant differences.

Some pixel-wise fingerprint segmentation methods compute only one feature at every pixel location, and threshold this feature to obtain the segmentation [22, 20, 28, 3]. Typically, the features computed are limited to filtering operators, or combination thereof, to limit the computational complexity. Obviously, the main drawback of using only one feature is that the results are directly tied to the expressiveness of the feature, and given the ad-hoc selection or construction of features in the literature, the results have tended to be unsatisfactory. Instead, and to avoid these limitations, other methods compute several features and then each pixel is classified independently either as part of the fingerprint or background [5, 40]. Using a classifier to combine features generally yields better results but the computation, to compute the input features and apply the classifier to each pixel, becomes high. For this reason, most approaches are block-wise based, commonly using blocks of 16×16 pixels. Block-wise approaches can have lower computation because the classifier is applied only once, and allow for more robust estimation of the features by averaging the features within each block. Moreover, because the classification is simple, the focus can shift towards computing better and more expressive features.

Generally, the key distinguishing characteristic between fingerprint segmentation methods is which set of features are used. A number of features have been proposed in the literature. Mehtre et al. [33] used information from the directional image to segment fingerprint images. The premise of this approach is that fingerprint regions have a well defined direction because of the ridge/valley pattern. In Mehtre et al. [32], this was augmented into a 'composite method' by also using the block variance, with the segmentation based on rule-based classification, empirically determined. Instead of block variance, variance computed along the direction orthogonal to the dominant ridge direction has also been utilized [38, 21]. Avoiding the explicit computation of the directional image, Bazen and Gerez [5, 4] proposed the use of coherence to determine whether local gradients are aligned along the same orientation, in addition to the local mean and variance. This feature set has also been used and extended in several other works [50, 39]. In particular, this feature set was extended in Zhao et al. [50] to include the contrast and mean energy ratio. The contrast is defined as the ratio between local variance and mean, and the mean energy ratio (MER) measures the ratio of energy in the main frequency

component with regard to the local non-DC average energy. The main idea is that the local frequency spectrum of a fingerprint has the energy concentrated onto two main peaks, and therefore the MER should be high. An alternative computation of the coherence was proposed in Wang et al. [44] using Gaussian-Hermite moments. The advantage of Gaussian-Hermite moments is that higher-order moments can express the presence of multiple parallel ridges. The limitation, however, is the need to carefully choose the scale parameter of the Gaussian-Hermite basis functions. Instead of using coherence, Chen et al. [10] proposed the use of the pixel clustering degree. The pixel clustering degree measures whether pixels with similar intensity values are grouped nearby, as occurs within a ridge/valley pattern, but does not depend on gradient direction. Also, Helfroush and Mohammadpour [17] proposed the dominant ridge score as an alternative to coherence. The dominant ridge score is the ratio of the number of pixels with the dominant gradient orientation, after quantization to 8 orientations, with regard to all the pixels in an image window. This idea is related to the use of the average gradient magnitude and variance of the gradient orientation proposed by Qi and Mei [36]. Other features proposed in the literature include the use the local binary pattern (LBP) operator [46, 45], and complementary variance energy [19].

Intuitively, the features proposed for fingerprint segmentation are conceptually similar in the sense that they attempt to quantify the observation that fingerprints are formed by a succession of alternating ridge/valleys with approximately constant orientation. Through these features, however, it is not entirely clear how the intrinsic structure of fingerprint images is being characterized, and therefore how these descriptors adapt to acceptable changes in the image. In contrast, the features proposed in this paper describe the manifold structure of fingerprint images, which remain invariant for a fingerprint region as long as the ridge/valley pattern is clearly discernible, even if this pattern changes because of contrast or noise.

The feature sets proposed by Bazen and Gerez [5, 4] and Zhao et al. [50] are used in the results section to establish a reference for evaluating the approach proposed here.

3. Fingerprint manifold characteristic features

In this section we analyze the manifold associated with fingerprint patches, and discuss how this knowledge can be utilized to derive features that characterize fingerprint regions. A brief review of manifold learning is included.

3.1. *Manifold learning*

Manifold learning is a research topic that has received significant attention in recent years, mainly fueled by developments in nonlinear dimensionality reduction techniques, such as ISOMAP [43], local linear embedding [42], and Laplacian eigenmaps [6], among others. Nevertheless, the concept has been around for quite some

time [16, 26] and is also the core idea behind linear dimensionality reduction methods, such as principal component analysis (PCA) [11].

The primary aim of manifold learning is to infer or characterize the structure of data by describing the manifold where the data lies. This means that, typically, high-dimensional data (or some high-dimensional representation as obtained in kernel methods or by embedding) actually lies in a relatively low-dimensional subspace (i.e., the manifold) that can be described by only a few parameters, or features. Indeed, for smooth manifolds with topology equivalent to the Euclidean topology (that is, the distance in the manifold is a smooth function of the Euclidean distance in the feature space), it can be shown that there exists an isomorphism between the space of features and the manifold [12]. Put differently, the dimensionality of the feature space is the number of parameters needed to parameterize the manifold, and the isomorphism characterizes its topology. Hence, all information about the data is expressed in this low-dimensional feature space, which can usually be related to some characteristic of interest.

If the data lies in a linear manifold (i.e., a linear low-dimensional subspace) of the high-dimensional data space, then principal component analysis (PCA) can be used to characterize the data structure [11]. PCA finds the orthonormal set of vectors $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_d\}$ that maximizes the variance of the projections of the data onto them. This means that projecting the high-dimensional data (say, of dimension D) onto those vectors yields a lower-dimensional representation (of dimension $d \leq D$) with a covariance matrix that is diagonal (i.e., uncorrelated projections) and with maximal entries. The vectors $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_d\}$ are called the *principal components* of the data. They are obtained by diagonalizing the covariance matrix of the data,

$$\mathbf{C} = \frac{1}{N} \sum_{i=1}^{N} (\mathbf{x}_i - \mu)(\mathbf{x}_i - \mu)^T, \tag{1}$$

where the $\mathbf{x}_i$'s, $i = 1, 2, \ldots, N$, are the N data points and μ denotes their mean. The principal components are given by the d eigenvectors of $\mathbf{C}$ corresponding to the d largest eigenvalues, which equal to the variance of the projections. The eigenvalues can also be used to estimate the manifold dimension d by guaranteeing that the eigenvalues corresponding to the principal components account for most of the data variance.

By deriving a low-dimensional characterization, manifold learning methods allow us to visualize high-dimensional data and understand the origin of its intrinsic data structure. Manifold learning is often used to find a lower dimensional representation of the data that is correlated with the perceived parameters. A typical example is to consider a set of images of an object taken at different rotations, and then attempt to verify the lower dimensional representation of the manifold formed by these images in terms of the rotation angle [43]. Similar ideas have also been applied, for example, to study sequences of images in movies [35], and to infer and constrain the space of shapes for cardiac MRI [48].
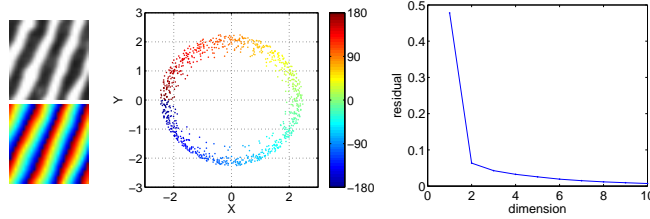
6   *Antonio R. C. Paiva and Tolga Tasdizen*

Fig. 2. PCA projection of 7×7 fingerprint patches in a small image. (Left) 30×30 window from which the patches were obtained. The phase image of the 2D PCA projections is shown in the bottom left. (Middle) Fingerprint patches projected onto the first two principal components and colored according to their phase; that it, each point is the 2D PCA projection of the 7×7 fingerprint patches. The points' colors match the colors of the corresponding center pixel as shown in the phase image in the bottom left. (Right) Residual variance versus the number of embedding dimensions.

3.2. *Manifold topology of fingerprint patches*

Unlike the example mentioned in the previous section, where a point corresponds to an image, for classification of image pixels (or groups of pixels) and analyzing the manifold structure arising from their interdependences, each pixel must have an associated representation. Instead of the value in a single pixel, image points can be embedded in a high-dimensional space by considering the pixel and its neighborhood; that is, by taking patches centered at the pixel location. Specifically, for an image I, an $M \times M$ patch corresponding to image point (x, y) is the ordered set of points $\{I(u, v) : |u - x| \leq (M - 1)/2 \wedge |v - y| \leq (M - 1)/2\}$. This embedding has the advantage that it preserves local and contextual information, since it completely characterizes the image joint distribution, up to the size of the neighborhood. Statistically, this data-driven representation corresponds to a sampling of the Markov Random Field model of the image.

Consider the patches extracted from a local region of a fingerprint, and their embedding to 2-D using, for example, principal component analysis (PCA) [11], as shown in Fig. 2. One can clearly observe the circular manifold formed by those fingerprint patches. (Obviously, nonlinear dimensionality reduction methods can be utilized for this analysis. However, because of their high computational complexity, they are impractical for fingerprint analysis and will not be considered further. Moreover, as discussed next, for image patches from local image windows, PCA suffices and is computationally simpler.) As asserted by the residual variance plot in Fig. 2(right),[a] most of the data variance is contained in the first two dimensions, as is to be expected since a circle is a 1-D topological structure in 2-D space. Since the points (i.e., fingerprint patches) form a circular manifold, they can be appropriately

[a]The residual variance with regards to embedding dimension k is the ratio of the error variance over total data variance when using k principal components for the embedding. Given the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_d$ of $\mathbf{C}$ (which correspond to the variance along the principal components), the residual variance with regards to dimension k is obtained as $r_k = \sum_{i=k+1}^{d} \lambda_i / \sum_{i=1}^{d} \lambda_i$.
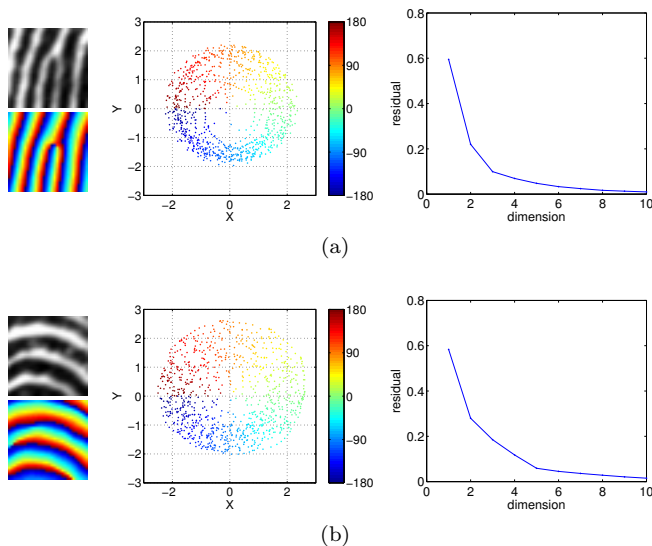
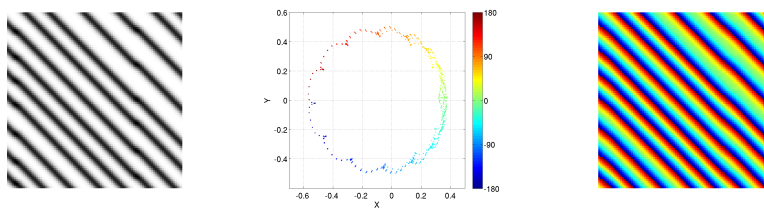Fig. 3. Similar to Fig. 2, but for images containing a minutia point (a) and near the fingerprint core (b).

Fig. 4. PCA projection of 7×7 patches from an image with parallel but *non-periodic* lines. (Left) Image from which the patches were obtained. (Middle) PCA projection of the patches projected onto 2-D, with color coded phase. (Right) Corresponding phase image.

characterized by their phase, which can be related back to the fingerprint region from where the patches were extracted. Note that the presence of minutiae in the image window and relatively small changes in orientation, as shown in Fig. 3, introduce additional variability in the linear projection of the manifold but the circular structure of the manifold does not change significantly. This is because the norm of the projections corresponds to the distance of the original fingerprint patches to the "mean patch", and thus captures primarily changes in contrast within a patch. Furthermore, Fig. 4 also shows that the manifold does not require the ridges to be parallel, unlike methods that "project" the image block along multiple directions [24, 33, 32].

It is insightful to understand why the manifold has this topology. If one considers a relatively small window of a fingerprint image (as the one shown in Fig. 2(top

left)), for most windows, we can observe a succession of ridges and valleys of approximately constant orientation. Hence, for these images, if one starts at a given point and marches through the image in the direction orthogonal to the ridges' orientation, one soon comes to a position in which the point and the context are very similar to those of the initial position. It is the repetition of 'context' that gives rise to the circular topology. The image window can thus be approximated as a wave, parameterized by a single argument θ corresponding to the phase of the wave. If the ridge and valley widths are equal, the wave pattern is a sinusoid; however, this does not need to be the case. Even if the ridge/valley pattern is aperiodic, i.e. not a sinusoid, the topology will still be that of a circle, although the distribution of phases will change (Fig. 4). In the case of fingerprint windows the manifold is actually a geometrical circle because the contrast is locally approximately constant (recall that contrast defines the norm of the centered embedded patches, and thus the radius of the circle). It must be remarked that we consider local windows such that the orientation is approximately constant and does not need to be considered explicitly as an argument in the parameterization. More specifically, we consider local image blocks because locality in the image space translates into locality in the embedding space, where the manifold is locally Euclidean [23, 41, 9]. In this way, we avoid the nonlinearity of the manifold when the orientation is considered, allowing for the use of PCA. On the other hand, if the orientation changes within an image block, the manifold becomes nonlinear introducing variability in the projection. This is because we are trying to characterize a nonlinear manifold using a linear method. This is noticeable in Fig. 3, where the nonlinearity of the manifold raises the relevance of additional principal components. Nevertheless, as long as changes in orientation are relatively small, the circular manifold structure will remain clearly distinguishable.

Unlike a conventional pattern recognition approach based on a static set of features, the approach just described has the advantage that the obtained features are found from and depend only on the data. Moreover, since the mapping from the data to features is found from a neighborhood of points, it incorporates characteristics of the whole neighborhood. This follows because the context information is included in the embedding of image points as patches and information about correlations with neighboring pixels is preserved. This approach also shares similarities to natural vision systems which extract information by contrasting a small region with its context. Still, the most remarkable property of this approach is that it allows us to focus on invariant descriptors on the data characteristics. Specifically, since we know that image patches from a fingerprint window form a geometrically circular manifold with near uniform phase distribution, we can simply measure the difference with regards to such a manifold structure. Such descriptors are invariant because they do not depend of the specific position or orientation of the sequence of ridges, provided the orientation is approximately constant, because the underlying
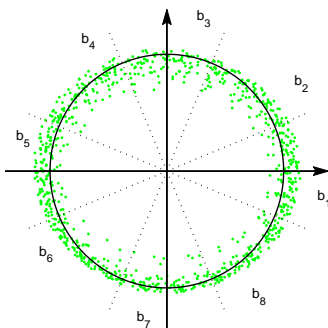
Fig. 5. Ratio and phase distribution features for fingerprint segmentation. The black circle has radius equal to the average norm and the division of the plane in phase bins (in this case, with 8 bins) are shown for the projection of the points in Fig. 2. The ratio r_i^- corresponds to the number of points inside the circle over the total number of points, and the values b_k, utilized in (2), are the ratio between the number of points within each phase bin to the total number of points.

manifold remains the same. Hence, if desired, it is possible to detect the fingerprint pattern even under low contrast, as will shown in the next section.

3.3. *Features for fingerprint analysis*

Knowledge of the manifold topology can be utilized for fingerprint segmentation since it characterizes the general distribution of the data embedded onto the feature space. Consequently, we can immediately suggest a set of features that best describes the data. This means that deviations from the typical fingerprint pattern can be easily captured since these points have a different manifold structure easily noticeable in their embedding and projection. These differences can then be utilized to distinguish a fingerprint region from background.

An obvious candidate feature would be to use an information-theoretic divergence, such as the Kullback-Leibler divergence [15], between the observed distribution and the distribution estimated from training. This approach, however, is cumbersome and computationally demanding since it would require the explicit estimation of the distribution. Given that the circular manifold formed by fingerprint patches is not too hard to characterize, we propose instead derived features that can be utilized to measure differences between the observed data manifold and a circle.

Two fundamental descriptors of the manifold geometry are the two leading eigenvalues of the PCA projection, denoted λ_{i1} and λ_{i2}, which correspond to the variance along the first two principal components. These eigenvalues characterize contrast and are related to the block variance, but provide additional information. For example, if the first eigenvalue is significantly larger than the second, then the image region primarily contains a gradient of intensities and is therefore unlikely to

be a fingerprint region.[b] Another important feature is the ratio of samples with a norm (distance from the origin in the projection) that is below the average norm, denoted r_i^- (see Fig. 5). For a circle, this value should be close to 0.5. This is an important descriptor since it allows us to easily discriminate from Gaussian noise, since in the latter case the value tends to be larger. (Note that the ratio of points with norm greater than the block mean norm could be utilized instead. Because the two ratios sum to one, they contain exactly the same information, and yield the same result.) Finally, one should check that the phase distribution is somewhat uniform, since this indicates the presence of the alternating ridge/valley pattern. This can be approximated by dividing the phase into a number of bins, say N_B, and computing the average difference magnitude to the expected value under a uniform distribution. For block i, we define this divergence as,

$$p_i = \sum_{k=1}^{N_B} \left| \frac{1}{N_B} - b_k \right|, \tag{2}$$

where b_k is the ratio between the number of data points with phase in the kth bin and the total number of data points (cf. Fig. 5). Summarizing, for the ith block, these features give rise to the feature vector,

$$\mathbf{x}_i = \left[\sqrt{\lambda_{i1}}, \sqrt{\lambda_{i2}}, r_i^-, p_i \right]^T. \tag{3}$$

The proposed features can be utilized by a classifier to segment fingerprints. If one does not have training data or intends to allow user interaction, then appropriately weighting and combining these features is not a trivial task. In these situations, it is best to have a single descriptor for which one only needs to set an appropriate threshold. For example, this could allow for the use of automatic threshold methods [44, 19]. Because we know the manifold topology, we can attempt to define a divergence measure specifically for this problem that bypasses the problems with an information-theoretic divergence as mentioned above.

As previously described, for fingerprint segmentation, a divergence is basically a measure that quantifies how much the data distribution differs from the circular manifold topology one expects for fingerprint regions. Let $\mathbf{y}_j^i$ denote the 2D PCA projection of the jth patch in the ith block, and let λ_{i1} be the corresponding largest eigenvalue of the PCA eigendecomposition. Then, a measure of divergence can be defined as

$$f_i = \frac{1}{\sqrt{\lambda_{i1}}} \left[\frac{1}{N} \sum_{j=1}^{N} \exp\left(-\frac{\left\| \mathbf{y}_j^i \right\|^2}{2\lambda_{i1}} \right) \right] \left[1 - \exp\left(-\frac{p_i}{\alpha} \right) \right], \tag{4}$$

---

[b]Note that this interpretation follows because these are eigenvalues of the covariance matrix and should not be confused with the eigenvalues of the structure tensor matrix used in Bazen and Gerez [4, 5].

where N is the number of patches per window, $\left\|\mathbf{y}_j^i\right\|$ is the Euclidean norm of the projected patch, and α is a user parameter. This divergence expresses various characteristics of the projected patches: variance, non-Gaussianity, constant radius, and uniform phase distribution. The variance of the projections onto the first principal component is λ_{i1}. Therefore, windows with good contrast have high variance which yields smaller values of f_i. The second term measures how well projected points are clustered around the origin, as one would get for windows containing unstructured values, such as Gaussian noise. For fingerprint windows, the projected points form a well defined circle on the tail of the exponential, and thus this term yields small values. Basically, these two terms correspond to a Gaussian density function, without the normalization by $\sqrt{2\pi}$. Approximate constant radius is enforced by the use of only the first eigenvalue in the second term rather than a covariance matrix. If the circle is distorted into an elongated ellipse, the second term yields higher values for points along the second principal component because they are closer to the origin (since $\lambda_{i2} \leq \lambda_{i1}$). The third term checks for a uniform distribution of points around the origin. This helps to exclude, for example, blocks from the edges of the fingerprint (which yield spurious minutiae) and single lines since, in these cases, the phase is skewed. The parameter α controls the allowable ratio of deviation from the uniform distribution. A value of $\alpha = 0.25/N_B$ seems reasonable for most cases, and was used for the results shown.

Either the divergence or features can be utilized for fingerprint segmentation. As mentioned, the use of the divergence has the advantage that one only needs to set the threshold, and is therefore better suited for user interaction. On the other hand, using the proposed feature vector has the advantage that the features can be optimally combined by the classifier to achieve the best performance for fingerprints captured using a specific condition, for example, the same sensor.

It must be noted that the segmentation obtained using either the proposed feature vector or the divergence will depend on contrast. This is because of the use of the projection eigenvalues, which corresponds to data variance, and thus contrast. This prevents our approach from being contrast invariant. However, in practice one needs a minimal contrast to ensure robust feature extraction and reject latent fingerprints in the background. To highlight the invariance of the topology with regards to contrast, consider the fingerprint image in Fig. 6a which contains a latent fingerprint in the background. If the term $1/\sqrt{\lambda_{i1}}$ is not included in the divergence, one obtains comparable values for the blocks in both the high-contrast fingerprint and the latent fingerprint in the background (Fig. 6c). Considering this term, as in (4), one introduces contrast dependency making it possible to distinguish the two regions, even though the divergence still yields relatively low values compared to other parts of the background (Fig. 6d).

(a) Fingerprint image.

(b) CLAHE contrast enhanced image.

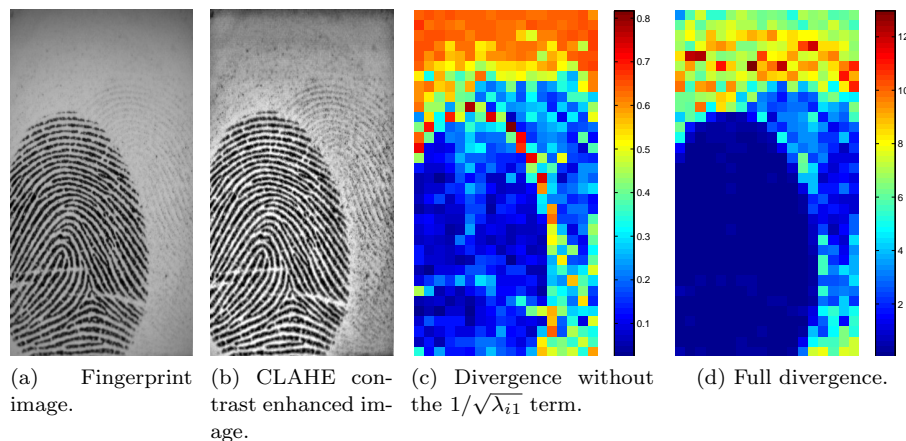(c) Divergence without the $1/\sqrt{\lambda_{i1}}$ term.

(d) Full divergence.

Fig. 6. Divergence of a fingerprint image with a latent fingerprint in the background. For ease of visualization, the contrast enhanced image is shown in (b). The divergence measures were computed for each 16×16 pixel block, using 7×7 patches, for the image shown in (a).
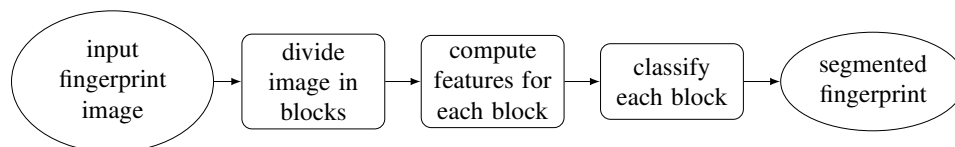

Fig. 7. Flowchart of the operations needed to segment a fingerprint image.

4. Fingerprint image segmentation

The features discussed in Sec. 3.3 can be utilized for fingerprint segmentation by classifying image blocks into fingerprint regions or image background. The sequence of operations to segment a fingerprint image is depicted in Fig. 7.

In this paper, we will utilize a linear classifier to classify each image block into a fingerprint block or background. More specifically, a Fisher linear discriminant is used because of its simple and consistent training, and because it has an analytical solution [13]. A perceptron linear classifier could also be utilized, as in Bazen and Gerez [5], but training requires the use of gradient decent, which depends on the learning rate, can get trapped in local minima, and the result can be unstable [13]. Clearly, nonparametric and/or nonlinear classifiers could also be used with likely benefits in performance [49]. The Fisher learning discriminant classifier is well suited for our purpose, however, because the main point of this analysis is to convey the better results due to the proposed improved features. Moreover, since most methods limit themselves to linear classifiers, this facilitates the comparison to earlier approaches. Finally, good performance regardless of the classifier is important in practice because AFIS systems typically have limited computational and memory

capabilities, and thus the choice of the classifier might depend on the available resources during system design.

The Fisher linear discriminant classifier finds a projection vector $\mathbf{w}$ such that the separation between two classes is maximized [13]. Put differently, the projection vector defines the hyperplane that best separates the two classes. Let Ω_0 and Ω_1 denote the background and fingerprint classes, respectively, and define the intraclass and inter-class scatter matrices, $\mathbf{S}_W$ and $\mathbf{S}_B$, as

$$
\begin{aligned}
\mathbf{S}_B &= (\mathbf{m}_1 - \mathbf{m}_0)(\mathbf{m}_1 - \mathbf{m}_0)^T, \\
\mathbf{S}_W &= \sum_{\mathbf{x} \in \Omega_1} (\mathbf{x} - \mathbf{m}_1)(\mathbf{x} - \mathbf{m}_1)^T \\
&\quad + \sum_{\mathbf{x} \in \Omega_0} (\mathbf{x} - \mathbf{m}_0)(\mathbf{x} - \mathbf{m}_0)^T,
\end{aligned}
\tag{5}
$$

where $\mathbf{x}$ denotes the vector of features, and $\mathbf{m}_1$, $\mathbf{m}_0$ are the mean feature vectors for each class. The projection vector is obtained as

$$
\mathbf{w} = (\mathbf{S}_W + \varepsilon \mathbf{I})^{-1}(\mathbf{m}_1 - \mathbf{m}_0),
\tag{6}
$$

where $\mathbf{I}$ is the identity matrix and ε is a small constant to ensure that $(\mathbf{S}_W + \varepsilon \mathbf{I})$ is invertible. Then, an image block is classified based on the projection value,

$$
v_i = \mathbf{w}^T \mathbf{x}_i,
\tag{7}
$$

using a decision threshold and direction (i.e., which, $v_i >$ threshold or $v_i <$ threshold, corresponds to fingerprint) chosen to minimize the error in the training set.

Note that situations in which the segmentation is achieved using only one feature, such as the divergence proposed earlier, are special cases. In these situations, learning the classifier simplifies to that of finding the decision direction and threshold.

5. Results

The proposed approach was tested on the datasets of the 2002 and 2004 Fingerprint Verification Competitions (FVC 2002 [1] and FVC 2004 [2]). A total of 8 datasets were included, corresponding to six different fingerprint scanners and two synthetic fingerprint datasets [29, 30, 31]. Each dataset contains 80 fingerprints (only the "part B" of the datasets was used), consisting of 8 fingerprints/subject from 10 subjects, captured under multiple conditions (e.g., normal, "dry" and "wet"). All fingerprints were manually segmented, with each 16×16 block being classified as fingerprint or background. A block was considered to be a fingerprint if the ridge/valley pattern was present and fully contained within the block. This means that blocks containing ridge terminations due to the boundary of the fingerprint area were not considered. This criteria was chosen because minutiae detection methods could easily confuse ridge terminations for minutia.

Segmentation results were computed using the features and the divergence proposed in Sec. 3.3 and the classifier detailed in Sec. 4. For comparison, segmentation

Table 1. Ratio of correctly segmented fingerprint image blocks using the MVC, xMVC, and the proposed feature set. The values are averages over 100 different pairs of randomly selected training/testing sets. The value within brackets corresponds to the worst segmentation accuracy in a fingerprint in the testing set, averaged over different training/testing sets. All differences in the results are statically significant (t-test, $p < 0.01$), except of course in between the proposed features and divergence in Dataset 4 of FVC 2002. The best result is shown in boldface.

		MVC	xMVC	Proposed features	Proposed divergence
FVC 2002	Dataset 1	0.907 [0.741]	0.944 [0.854]	**0.972** [0.895]	0.960 [**0.897**]
	Dataset 2	0.916 [0.755]	0.917 [0.747]	**0.956** [**0.847**]	0.948 [0.823]
	Dataset 3	0.912 [0.763]	0.911 [0.764]	**0.929** [**0.774**]	0.923 [0.739]
	Dataset 4	0.925 [0.731]	0.913 [0.719]	**0.960** [**0.881**]	**0.960** [0.877]
FVC 2004	Dataset 1	0.952 [0.884]	0.957 [0.896]	**0.973** [**0.937**]	0.958 [0.895]
	Dataset 2	0.834 [0.501]	0.848 [0.541]	**0.934** [**0.755**]	0.869 [0.585]
	Dataset 3	**0.964** [0.866]	0.963 [0.867]	0.961 [**0.874**]	0.957 [0.833]
	Dataset 4	0.917 [0.718]	0.891 [0.599]	**0.950** [**0.814**]	0.945 [0.804]

results were computed using the feature set proposed by Bazen and Gerez [5], which consists of the block mean, variance and coherence. We shall refer to this as the "MVC feature set." In addition, we computed segmentation results using the extended MVC features set used in Zhao et al. [50], which further includes the block contrast and mean energy ratio (MER) (see Sec. 2 for details), henceforth denoted as the "xMVC feature set." In all experiments, the Fisher linear discriminant classifier described in Sec. 4 was utilized.

The results for the segmentation accuracy (without post-processing) are presented in Table 1. The blocks were 16×16 pixels, and 7×7 image patches were used for the embedding in the computation of the proposed features and divergence. These results show that the proposed feature set yields a consistently better segmentation. The improvement in overall segmentation accuracy ranges from 1.6% to 8.4%, except for dataset 3 of FVC 2004 in which using the proposed features performs 0.3% worse. From analyzing the segmented fingerprint images, it seems that the classifier using the MVC feature set marked these blocks as fingerprint due to their high variance, even though they have no discernible fingerprint pattern. Using the proposed features, such blocks were marked as background precisely because of the absence of the ridge/valley pattern. Such image blocks were labeled as fingerprint in the manual segmentation because they arise isolated in the middle of the fingerprint, but this choice is arguable. The relevance of the proposed features becomes even more evident when considering the worst segmentation accuracy in a testing fingerprint, in which case the difference can be as high as 21.4% (dataset 2 of FVC 2004). This demonstrates that using the proposed features yields much more consistent segmentation results.

The results using the proposed divergence are slightly worse that using the proposed features, but still better overall than using either the MVC or xMVC feature
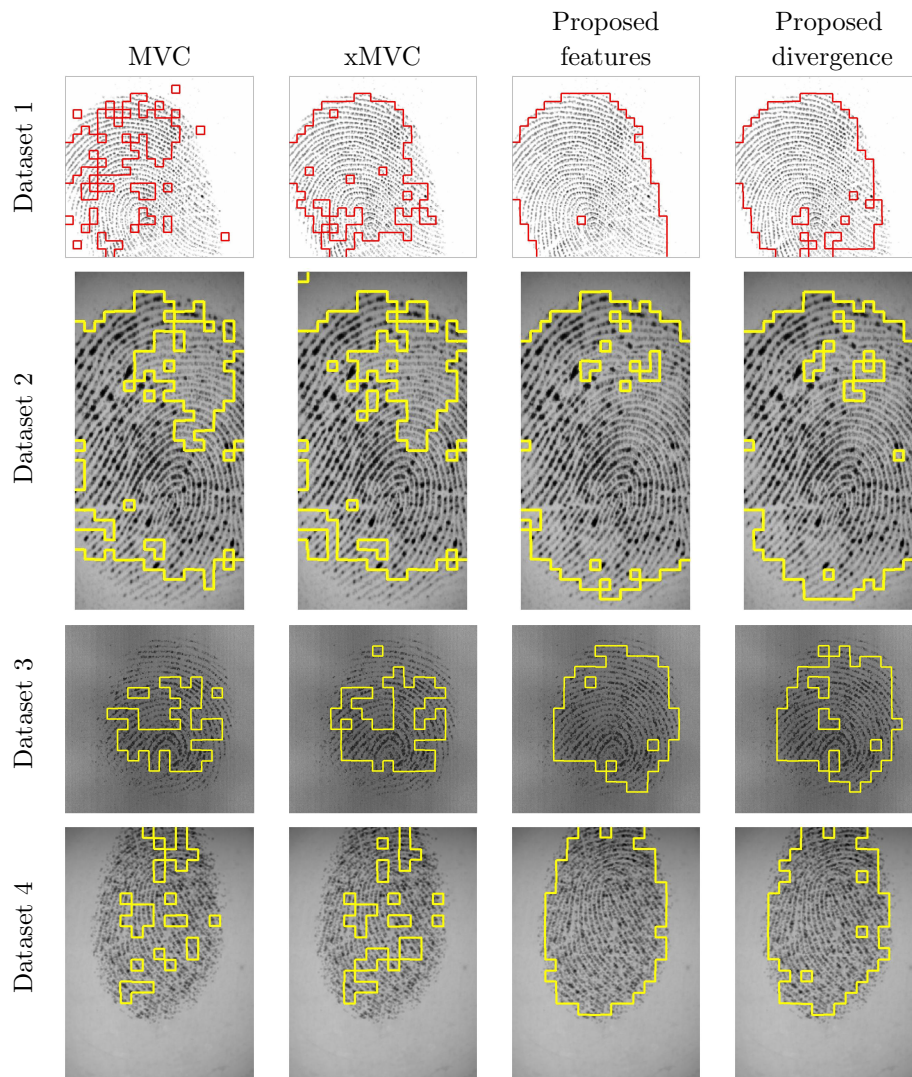
Fig. 8. Segmentation results for fingerprint images from FVC 2002.

sets. The fact that the divergence, which nonlinearly combines several elementary features, did not surpass the performance of an optimal linear classifier with the proposed features was not surprising. This is due to the ability of the latter to fine tune the weights for each dataset. On the other hand, this suggests that the segmentation accuracy using the proposed features can be further improved using a nonlinear classifier.

The better behavior of the proposed features and divergence can be verified visually in the segmented fingerprint images shown in Figs. 8 and 9. It can be
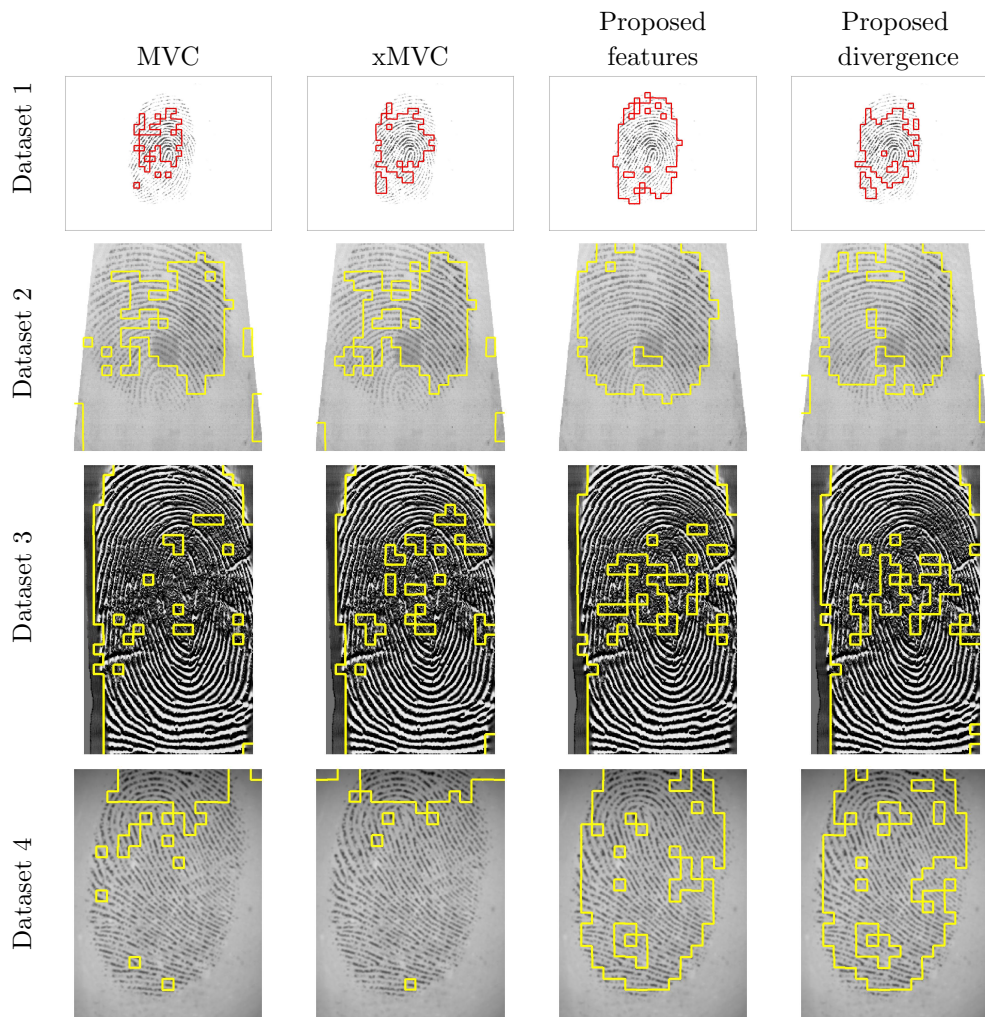
16   *Antonio R. C. Paiva and Tolga Tasdizen*

Fig. 9. Segmentation results for fingerprint images from FVC 2004.

seen that although the segmentation using the MVC and extended MVC feature sets is successful for relatively high-quality / high-contrast fingerprint regions, its performance drops significantly if these characteristics change, yielding very poor segmentations for some fingerprint images. In contrast, the proposed features and divergence show a much more robust and consistent behavior. This is due to the use of the manifold topology, which remains noticeable even in the presence of faint lines, as explained in Sec. 3.3 (see Fig. 6). This robustness is further observed in Fig. 10 which showcases the worst segmentation result using the proposed features. Although the results are clearly unsatisfactory, the segmentation using the proposed

October 18, 2012   22:57

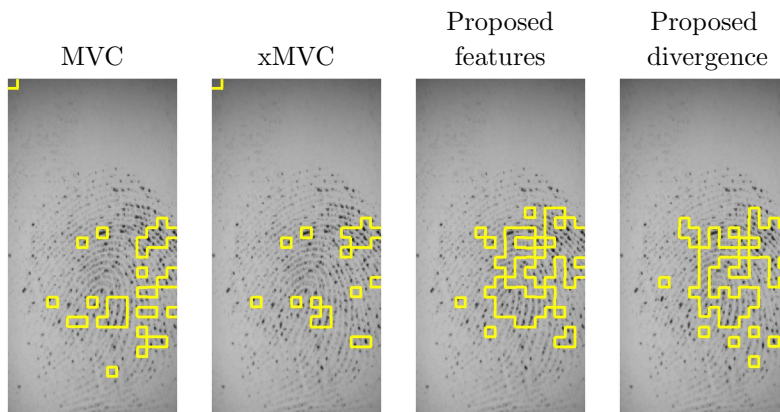| MVC | xMVC | Proposed features | Proposed divergence |

Fig. 10. Fingerprint image for which we observed the worst segmentation results using the proposed features (from FVC 2002, Dataset 2).

features or divergence is still reasonable (compare with the MVC or xMVC results in Dataset 4 shown in Fig. 9) and much better than the result obtained using the MVC or xMVC feature sets.

It must be remarked that these results could be improved further through pre- or post-processing, or by using a different classification strategy. Pre-processing could be used to remove some noise and potentially mitigate variations in contrast between images. Likewise, post-processing could be used to compensate for minor errors in the block segmentation due to image noise, variability, and limitations of the classifier. Although, for the purpose of this paper, post-processing was not used, the results clearly suggest that, if post-processing is employed, the errors obtained with the proposed approach would seem much more manageable. The results could be further improved by providing block context to the classifier (e.g., the classifier observes the features from its block and neighboring blocks) for example. Moreover, utilizing a nonlinear classifier (e.g., AdaBoost or SVM [8]) might be able to amplify the context of the features to the point of compensating for extreme variations in contrast, such as those shown in Fig. 10. Additionally, these improvements to the classifier may be suplemented with approaches such as adaptive thresholding [14].

6. Discussion

At first glance, the computational complexity of the proposed features seems much higher than that required to compute the MVC or extended MVC features. Indeed, in our Matlab implementation, computing the proposed features required about 4 to 5 times more CPU time in single-thread mode than the features with which we compared (cf. Table 2). However, while computation of the comparison features uses inherently optimized routines, such as filtering and fast Fourier transforms, the computation of our features does not. Although intricate, it is relatively easy to

reduce the computational complexity for the calculation of the covariance matrix of image neighborhoods and its eigendecomposition, which are the two most time consuming operations in the computation of the proposed features and divergence. Because the same spatial relationship is encountered between different pixels within an image neighborhood, most elements of the covariance matrix are repeated. As an example, note that the diagonal elements of the covariance matrix are all equal and thus need to be computed only once. This is similar to the Toeplitz structure of covariance matrices of vectors from 1-D data but, for image patches, one has to further handle the relationships between pixels in the edges of the patches. It can be easily shown that, for $n \times n$ image patches, the number of unique entries of the covariance matrix that need to be computed is only $n^2 + (n-1)^2$, compared to n^4 in a naive implementation. This represents a reduction in computation complexity greater than $n^2/2$. For 7×7 patches, the savings in computational complexity would be about $28 \times$. Moreover, the computational complexity of the eigendecomposition can also be significantly reduced because it is known before hand that only the first two eigenvectors are needed in this case.

In the results shown, we tacitly considered 16×16 image blocks and 7×7 image patches. Clearly, the performance of our approach depends on these settings since, in a sense, these parameters define the scale one uses to interpret the image. However, we have found empirically their choice not to be critical (cf. Table 3). Generally speaking, the block size should be big enough such that the alternating ridge/valley pattern can be observed, while small enough to avoid having major changes in the orientation of the ridges, which introduces nonlinearities in the manifold structure. Hence, for a fingerprint image of typical resolution, block sizes between 16×16 to 24×24 pixels should yield similar results. Likewise, the patch size represents a trade-off, albeit between context information and computational complexity. On one hand, the patch size should be large to include as much context information as possible in the embedding. On the other hand, the patch size should be small enough to avoid computation of large covariance matrices and their eigendecomposition. In our experiments, we found that 7×7 image patches presents a good trade-off. For example, using 9×9 image patches the segmentation accuracy only improved at most 0.5%. However, it more than doubled the computation time.

7. Conclusion

This paper describes an alternative approach to fingerprint segmentation. The core idea is the use of the manifold topology of fingerprint image patches as a descriptor of the intrinsic structure of fingerprint images. Indeed, it is shown that an embedding of fingerprint image "points" in a high-dimensional space, in the form of image patches, forms a simple and highly regular circular manifold. This perspective is very useful since this insight allows us to naturally derive or suggest a number of features that characterize the intrinsic properties of a fingerprint. Hence, in this approach, fingerprint segmentation can be formulated as the problem of classify-

ing a fingerprint region (e.g., image block) based on a deviation from the expected topology.

The improved performance and robustness was shown in eight datasets from FVC 2002 and FVC 2004. In spite of the diverse characteristics of the fingerprint images in terms of contrast, noise and conditions (e.g., "wet" vs. "dry") across and within each dataset, the approach presented yielded more a robust and consistent segmentation. This is very important in actual applications where the capture conditions may vary tremendously. The robustness follows from the use of the manifold topology, which is invariant to contrast, orientation and robust to noise as long as the alternating ridge/valley pattern is noticeable.

From a different perspective, it is interesting to verify that utilizing the manifold structure to infer information about the image shares similarities to visual systems. This is because visual systems are capable of extracting salient features just by contrasting these points with their local context and/or with the perception of the image structure at a broader level [27]. Roughly speaking, this is precisely what is obtained by embedding the image points and locally computing the manifold topology.

The knowledge and use of the manifold topology was utilized here to characterize fingerprint images for segmentation, but the same concept may have potential implications in many other problems, such as texture classification and segmentation, for example. By characterizing and studying the manifold topology of embedded points from images, one gains an understanding of the images' characteristic structure [34]. Given a problem, this knowledge can then be utilized to naturally suggest or derive descriptors that best characterize the structure of the image for that problem, as shown here.

Appendix

Table 2. Average computation time (in seconds) to compute the features for each image on an Intel Xeon X7350 at 2.93GHz with Matlab restricted to a single-thread.

		MVC	xMVC	Proposed features	Proposed divergence
FVC 2002	Dataset 1	0.162	0.223	0.606	0.545
	Dataset 2	0.183	0.256	0.848	0.780
	Dataset 3	0.102	0.143	0.498	0.461
	Dataset 4	0.124	0.173	0.576	0.532
FVC 2004	Dataset 1	0.364	0.493	1.114	0.976
	Dataset 2	0.138	0.193	0.636	0.585
	Dataset 3	0.155	0.216	0.739	0.682
	Dataset 4	0.124	0.173	0.561	0.516

Table 3. Comparison of segmentation results using the proposed feature set or divergences as a function of the image patch size. The values are averages over 100 different pairs of randomly selected training/testing sets. The value within brackets corresponds to the worst segmentation accuracy in a fingerprint in the testing set, averaged over different training/testing sets.

		Proposed features			Proposed divergence		
		5×5	7×7	9×9	5×5	7×7	9×9
FVC 2002	Dataset 1	0.972 [0.893]	0.972 [0.895]	0.969 [0.888]	0.956 [0.878]	0.960 [0.897]	0.963 [0.905]
	Dataset 2	0.954 [0.854]	0.956 [0.847]	0.952 [0.841]	0.946 [0.832]	0.948 [0.823]	0.949 [0.829]
	Dataset 3	0.932 [0.769]	0.929 [0.774]	0.934 [0.799]	0.909 [0.677]	0.923 [0.739]	0.935 [0.794]
	Dataset 4	0.959 [0.866]	0.960 [0.881]	0.961 [0.885]	0.955 [0.850]	0.960 [0.877]	0.960 [0.897]
FVC 2004	Dataset 1	0.975 [0.937]	0.973 [0.937]	0.971 [0.932]	0.958 [0.903]	0.958 [0.895]	0.960 [0.891]
	Dataset 2	0.937 [0.766]	0.934 [0.755]	0.933 [0.755]	0.857 [0.537]	0.869 [0.585]	0.880 [0.654]
	Dataset 3	0.960 [0.872]	0.961 [0.874]	0.958 [0.880]	0.956 [0.810]	0.957 [0.833]	0.958 [0.857]
	Dataset 4	0.951 [0.805]	0.950 [0.814]	0.951 [0.824]	0.940 [0.780]	0.945 [0.804]	0.949 [0.828]

References

1. Fingerprint Verification Competition 2002, FVC 2002. http://bias.csr.unibo.it/fvc2002/.

2. Fingerprint Verification Competition 2004, FVC 2004. http://bias.csr.unibo.it/fvc2004/.

3. A. Baig, A. Bouridane, and F. Kurugollu. A corner strength based fingerprint segmentation algorithm with dynamic thresholding. In *Proc. Intl. Conf. on Pattern Recognition*, Dec. 2008.

4. A. M. Bazen and S. H. Gerez. Directional field computation for fingerprints based on the principal component analysis of local gradients. In *ProRISC Workshop on Circuits, Systems and Signal Processing*, Veldhoven, Netherlands, Nov. 2000.

5. A. M. Bazen and S. H. Gerez. Segmentation of fingerprint images. In *ProRISC Workshop on Circuits, Systems and Signal Processing*, Veldhoven, Netherlands, Nov. 2001.

6. M. Belkin and P. Niyogi. Laplacian eigenmaps for dimensionality reduction and data representation. *Neural Comp.*, 15:1373–1396, 2003.

7. S. D. Bharkad and M. Kokare. Performance evaluation of distance metrics: application to fingerprint recognition. *Int. J. of Pat. Rec. and Artificial Intel.*, 25(6):777–806, 2011.

8. C. M. Bishop. *Pattern Recognition and Machine Learning.* Springer, 2006.

9. M. Brand. Charting a manifold. In *Advances Neural Information Processing Systems*, Vancouver, BC, Canada, Dec. 2002.

10. X. Chen, J. Tian, J. Cheng, and X. Yang. Segmentation of fingerprint images using linear classifier. *J. Appl. Sig. Proc.*, 2004(4):480–494, 2004.

11. K. I. Diamantaras and S. Y. Kung. *Principal Component Neural Networks: Theory and Applications.* John Wiley & Sons, 1996.

12. D. L. Donoho and C. Grimes. Image manifolds which are isometric to euclidean space. *J. Math. Imag. Vis.*, 23(1):5–24, July 2005.

13. R. O. Duda, P. E. Hart, and D. G. Stork. *Pattern classification.* Wiley Interscience, 2nd edition, 2000.

14. R. C. Gonzalez and P. Wintz. *Digital image processing.* Addison Wesley, 2nd edition, 1987.

15. R. M. Gray. *Entropy and Information Theory.* Springer-Verlag, New York, NY, USA, 1990.

16. T. Hastie and W. Stuetzle. Principal curves. *J. Am. Stat. Assoc.*, 84(406):502–516, June 1989.

17. M. S. Helfroush and M. Mohammadpour. Fingerprint segmentation. *World Appl. Sci. J.*, 6(3):303–308, 2009.

18. L. Hong, Y. Wan, and A. Jain. Fingerprint image enhancement: Algorithm and performance evaluation. *IEEE Trans. Pattern Anal. Machine Intell.*, 20(8):777–789, Aug. 1998.

22   *REFERENCES*

19. Z. Hou, W. Yau, N. Than, and W. Tang. Complementary variance energy for fingerprint segmentation. In *Proc. Advances in Multimedia Modeling, LNCS 4903/2008*, pages 113–122, 2008.

20. C. Hu, J. Yin, E. Zhu, H. Chen, and Y. Li. A composite fingerprint segmentation based on Log-Gabor filter and orientation reliability. In *Proc. IEEE Intl. Conf. on Image Processing*, pages 3097–3100, Sept. 2010.

21. A. Jain, L. Hong, and R. Bolle. On-line fingerprint verification. *IEEE Trans. Pattern Anal. Machine Intell.*, 19(4):302–314, Apr. 1997.

22. H. Jia and P. Tang. An improved fingerprint image segmentation algorithm. In *Proc. Intl. Conf. on Comp. Sci. and Electronics Eng.*, pages 118–120, Mar. 2012.

23. N. Kambhatla and T. K. Leen. Dimension reduction by local principal component analysis. *Neural Comp.*, 9:1493–1516, 1997.

24. S. Karimi-Ashtiani and C.-C. J. Kuo. A robust technique for latent fingerprint image segmentation and enhancement. In *Proc. IEEE Intl. Conf. on Image Processing*, pages 1492–1495, 2008.

25. P. W. Kwan, J. Gao, Y. Guo, and K. Kameyama. A learning framework for adaptive fingerprint identification using relevance feedback. *Int. J. of Pat. Rec. and Artificial Intel.*, 24(1):15–38, 2010.

26. M. LeBlanc and R. Tibshirani. Adaptive principal surfaces. *J. Am. Stat. Assoc.*, 89(425):53–64, Mar. 1994.

27. R. Linsker. Self-organization in a perceptual network. *Computer*, 21(3):105–117, 1988.

28. H. Liu and L. Wang. Fingerprint segmentation based on Harris corner strength. *Comp. Eng. and App.*, 47(15):206–208, May 2011.

29. D. Maio, D. Maltoni, R. Cappelli, J. L. Wayman, and A. K. Jain. FVC 2002: Second fingerprint verification competition. In *Proc. Intl. Conf. on Pattern Recognition*, 2002.

30. D. Maio, D. Maltoni, R. Cappelli, J. L. Wayman, and A. K. Jain. FVC 2004: Third fingerprint verification competition. In *Proc. Intl. Conf. on Biometric Authentication*, 2004.

31. D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar. *Handbook of Fingerprint Recognition*. Springer, New York, NY, 2003.

32. B. M. Mehtre and B. Chatterjee. Segmentation of fingerprint images – a composite method. *Pat. Rec.*, 22(4):381–385, 1989.

33. B. M. Mehtre, N. N. Murthy, S. Kapoor, and B. Chatterjee. Segmentation of fingerprint images using the directional image. *Pat. Rec.*, 20(4):429–435, 1987.

34. A. R. C. Paiva and T. Tasdizen. Detection of salient image points using manifold structure. In *Proc. Intl. Conf. on Pattern Recognition, ICPR-2010*, Istanbul, Turkey, Aug. 2010.

35. R. Pless. Using ISOMAP to explore video sequences. In *Proc. Intl. Conf. on Computer Vision*, pages 1433–1440, 2003.

36. J. Qi and M. Xie. Segmentation of fingerprint images using the gradient vector field. In *Proc. IEEE Conf. on Cybernetics and Intelligent Systems*, pages 543–545, Sept. 2008.

37. N. Ratha and R. Bolle, editors. *Automatic Fingerprint Recognition Systems.* Springer, New York, NY, 2003.

38. N. K. Ratha, S. Chen, and A. K. Jain. Adaptive flow orientation-based feature extraction in fingerprint images. *Pat. Rec.*, 28(11):1657–1672, 1995.

39. C. Ren, Y. Yin, J. Ma, and G. Yang. A linear hybrid classifier for fingerprint segmentation. In *Proc. Intl. Conf. on Natural Computation*, pages 33–37, Oct. 2008.

40. D. A. Rojas and J. L. Aching S. An improved method for segmentation of fingerprint images. In *Proc. Electronics, Robotics and Automotive Mechanics Conf.*, pages 20–23, 2006.

41. S. Roweis, L. K. Saul, and G. E. Hinton. Global coordination of local linear models. In *Advances Neural Information Processing Systems*, pages 889–896, Vancouver, BC, Canada, Dec. 2001.

42. S. T. Roweis and L. K. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290:2323–2326, Dec. 2000.

43. J. B. Tenenbaum, V. de Silva, and J. C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290:2319–2323, Dec. 2000.

44. L. Wang, H. Suo, and M. Dai. Fingerprint image segmentation based on Gaussian-Hermite moments. In *LNCS 3584 – Advanced Data Mining and Applications*, pages 446–454, 2005.

45. C. Wen, T. Guo, and Y. Zhou. A novel and efficient algorithm for segmentation of fingerprint image based on LBP operator. In *Proc. Intl. Conf. on Information Technology and Computer Science*, pages 200–204, 2009.

46. B. Zhang, J. Zhang, H. Jiang, and X. Jing. A new algorithm of incomplete fingerprint segmentation based on LBP. In *Proc. Intl. Conf. on Cloud Comp. and Intel. Sys.*, pages 610–613, Sept. 2011.

47. H. Zhang, D. Miao, and C. Zhong. Modified principal curves based fingerprint minutiae extraction and pseudo minutiae detection. *Int. J. of Pat. Rec. and Artificial Intel.*, 25(8):1243–1260, 2011.

48. Q. Zhang, R. Souvenir, and R. Pless. On manifold structure of cardiac MRI data: Application to segmentation. In *Proc. IEEE Conf. on Computer Vision and Pattern Recognition*, pages I:1092–1098, 2006.

49. S. Zhao, B. Fan, Z. Zhao, and W. Tao. Fingerprint image segmentation method based on normal distribution model. In *Proc. Chinese Control and Decision Conf.*, pages 1753–1756, Sept. 2008.

50. S. Zhao, X. Hao, and X. Li. Segmentation of fingerprint images using support vector machines. In *Proc. Intl. Symp. on Intelligent Information Technology Application*, pages 423–427, Dec. 2008.